

Comments on ECOFRAM Draft Aquatic Report

R. Graney
Bayer Corporation

Overall, I was extremely impressed with the depth and detail in the report. Obviously a lot of time and effort went into the discussions and report preparation.. There is no doubt that this report represents an important reference material for those performing Aquatic Risk Assessments. However, in order for this report to be most effective as a tool for change of the current risk assessment paradigm, considerable effort is still necessary in focusing and prioritizing the many recommendations provide.

Overall, my comments are arranged in the following order:

1. Response to specific questions posed to the Reviewers,
2. General comments,
3. Detailed, page specific comments.

Response to specific questions posed to the Reviewers

ECOFRAM Workshop Panel Members were specifically asked to address 5 questions. Below are my responses to these questions:

1. Is the draft report scientifically sound?

The framework, concepts and theories presented within this report are all scientifically sound. Although “scientific soundness” is a critical question, it is not the most important question. To me, a more important question is .. can this framework and associated concepts be implemented within the current regulatory paradigm? With this question, I have greater concerns, many of which are outlined below.

2. Did the ECOFRAM Workgroup address the “Charge to the Terrestrial and Aquatic Workgroups” identified in the background document? To address this question I will paraphrase various portions of the background document.
- *“To develop and validate risk assessment tools and processes that address increasing levels of biological organization (eg. individuals, populations, communities, ecosystems), accounting for direct and indirect effects that pesticides may cause.....work groups will first address direct acute and chronic effects of pesticides on individuals and populations of high risk species. Work groups are charged with developing a process and*

tools for predicting the magnitude and probabilities of adverse effects to nontarget aquatic and terrestrial species.”

Relative to this charge, ECOFRAM has provided an excellent outline and overview of a risk assessment process which should greatly improve the current deterministic, risk quotient approach. However there was certainly no “validation” of the risk assessment tool and, in reality, I am not sure how ECOFRAM could have validated the process. Before implementation, it is critical that this process be applied to a variety of case studies to determine exactly how it may be used under “real world” regulatory circumstances. Also, based on the comments made by the SAP and the Charge to the Work Groups, it seems that the primary focus of the new process was to improve the probabilistic nature of the assessments. There is no doubt that this is accomplished; however it also apparent that there are numerous recommendations and components discussed within the document which are not directly linked with probabilistic assessments.

- *“The tools that are developed need to have reasonable scientific certainty and be capable of acceptable validation within a reasonable time frame”*

Most of the tools recommended are reasonable; however it is doubtful that many of the proposed exposure models can be validated within a reasonable time frame.

- *“Probabilistic techniques developed should use existing fate and effects data where possible. However, in developing new methodologies and improving risk estimates, it may be necessary to modify or discontinue current tests or to develop new ones”*

As discussed in greater detail below, a plethora of potential new fate and effects tests are recommended; however clear guidance is never provided for when such studies should be required or no procedure is recommended, such as sensitivity analysis, to determine if such studies would provide comparative value to the risk assessment. In addition, there was never any discussion concerning what current studies do not add value to the newly proposed risk assessment scheme and thus should be dropped.

- *“Methods should be specific enough to allow different risk assessors supplied with the same information to estimate similar values of risk.”*

This specificity certainly exists at Tier I and, assuming standard modeling scenarios are developed as proposed for Tier II, it is likely that consistency in risk estimations can be attained at this Tier. However once entering Tiers III and IV, consistency in the estimation of risk will be difficult. Given the flexibility in the design this is not surprising and it will be a difficult challenge for EPA management to ensure that consistency is maintained.

3. What are the limitations for predicting risk using the approach described in the draft report.

Three overall limitations include: 1) availability of adequate data, 2) actually developing and maintaining the required exposure models and 3) training of personnel. The complexity of the upper Tier assessments is extensive and having adequately trained personnel to conduct and/or evaluate such assessments will be very difficult. Greater detail addressing concerns with the proposed approach are provided below.

4.What areas of the report need to be strengthened??...

Overall the report was extremely thorough and informative. Critical areas to be addressed include: (1) The executive summary will be very important. The report provides so much detail and so many recommendations that summarizing these into a concise and readable executive summary will be difficult, but essential. (2) Much greater effort is necessary in prioritizing recommended actions, especially in the Exposure Chapter. Everything can not be done immediately. It is essential for the Group to identify the most important factor for which the greatest “value” can be attained in the shortest time frame. (3) The report is extremely redundant in some spots. If possible, this redundancy should be reduced - primarily between Chapter 2 and Chapters 3 and 4. (4) There should be a common chapter which attempts to standardize terminology and concepts between the aquatic and terrestrial reports. Other suggested improvements can be found below.

5. At what point in the risk assessment process is the certainty level high enough to support the consideration of risk mitigation? What is the minimum level of technical information and scientific understanding that is necessary to evaluate whether risk mitigation would be necessary and / or effective?

I find the first question a little bit confusing since I consider mitigation an ongoing evaluation which naturally occurs during all phases of the risk assessment process. Once a Tier I assessment is complete and a specific unacceptable risk is identified, then the process for considering exposure reduction options begins. Given this, I would think that the minimum information necessary would be a least some PRZM / EXAMS runs. This would allow some minimum sensitivity analysis on potential mitigation options. Obviously exactly where and how mitigation is actually integrated will be case specific and depend upon the magnitude and extent of the potential problem.

General comments on the Aquatics Report

There is an extensive array of recommendations / conclusions with little indication of the priorities for implementation. Many of the points raised or made can be implemented quickly and relatively soon; whereas others are a long way off. The report should outline what can be done in the near

term and how to accomplish this and what should be considered longer term objectives.

Training and available expertise to implement the recommended changes is a major area of concern. It is unclear who is expected to conduct these higher Tier assessments - the registrant or EPA? If industry is expected to conduct these, then clear and concise guidance will be required. If EPA plans to conduct the higher Tier assessments, then I would question the availability of adequate resources. I would recommend that such assessments be conducted by industry and reviewed by EPA (as study reports are reviewed now).

The report should include a discussion on data quality. Not all data are equal, and a weighting factor should be considered for evaluating available data on a product, and using these data in a risk assessment. A data quality evaluation may also provide some insight on data points that are outliers.

Case studies must be developed and worked through to actually evaluate the usefulness and effectiveness of the proposed changes. I would question setting up a new, potentially much more expensive and complicated process without evaluating whether the proposed changes truly add value (reduce uncertainty and thus allow for better or easier decisions). It is unclear to me whether it is expected that most products are going to need Tier II or higher assessments. I get the impression from the document that the ECOFRAM participants believed that higher Tier assessments will often be required. If this is the case, what is the cost of all this additional work??

From another perspective, is there the belief that the current RQ system is inadequate and letting numerous products with unacceptable risk “slip through the cracks” and that these will be “discovered” by this new risk assessment process? If this is not the case, then does this new process simply provide a better approach for evaluating those products presenting high risk?If this new process is only to be applied to the few (??) products which do not pass Tier 2, then how much effort should really go into developing new tests and guidelines?

I have never completely understood the differences in the number of test species required for fish, invertebrates and plants / algae. If one considers both marine and freshwater species, tests are required on 3 fish species, 3 invertebrates and 5 plants. I understand the reason given for testing more plants is the wider range in susceptibility (is there a reference for this?); however it does not match closely with the historical assessment endpoints used by the EPA. ECOFRAM should address whether having the additional plant species adds greatly to the proposed probabilistic risk assessment scheme and how this fits with EPA’s history of not including aquatic plants as one of their key assessment endpoints.

Role of the risk managers in determining the need for additional data needs to be clearly articulated. The need for additional data is based upon the uncertainty in the risk assessment and the acceptable level of uncertainty is partially determined by the risk managers. This point needs to be more clearly outlined.

The open-ended “tool box” represents an area of concern. Scientists, whether in government, academia or industry, always want more data! The need for more data is based upon the uncertainty in the risk assessment; **however the “measurement” of acceptable uncertainty is not defined?? It is essential that clear guidance be provided by ECOFRAM in determining and presenting this uncertainty such that consistent decisions can be made regarding the need for additional data.** It is possible that “sensitivity analysis” could be used to determine the need for additional data. In any case, there is always uncertainty and it is this uncertainty which makes decision making difficult. In large institutions, uncertainty can be used as a reason for not making a decision under the guise of needing more information. It is critical that the Risk Managers take a strong role in forcing decisions in the face of uncertainty.

Relative to the decision of when to require additional studies, it is essential that before any new studies become new guideline requirements, cost / benefit analysis be performed. This could be included in the case studies which will hopefully be conducted prior to implementation of any changes. In these case studies, the relative value added to the risk assessment could be compared to the cost of conducting the studies.

How is acceptable versus unacceptable risk defined?? It is recommended that it still be defined at Tier I based on quotient method, but what about the upper tiers? How and when will it be defined? From a risk management perspective, acceptable risk will vary from product to product based upon alternatives available etc., but some guidance is still required.

The assessment endpoints need to be defined. I did not find any discussion on this factor; however when one reads EPA's various risk assessment framework documents, it is stated that this is an essential component which is critical to define early in the process. Since the assessment endpoints should be the same for all products, or classes of products, these should all be defined up front. It was stated in the report that industry should be able to move through Tier III without significant interaction or feedback from the Agency. If this is the case, then a clear definition of the assessment endpoints is required. Relative to this, it is stated in the report (page 2-23) that ... “except in the special case of protected species, the environmental entity to be protected - the assessment endpoint - is not the individual but the population”. I agree with this statement. If possible, this, and other relevant statements addressing the appropriate assessment endpoints should be outlined by ECOFRAM. If not, EPA must address this issue before any changes in the process are considered.

The report seems to have relatively little discussion about more site-specific evaluations when lower tier analyses suggest an aquatic risk issue exists, except at the highest tier by doing more extensive exposure (i.e., GIS) and effects work. Considerations of specific crop/compound combinations can be incorporated earlier in the process, especially determining where a product is primarily used. Minor uses will often not justify extensive higher Tier analyses and will need to be handled earlier in the process in a more simplistic fashion.

Statements regarding chronic risk assessments and comparison to instantaneous test concentrations must be clarified. For chronic assessments, comparing chronic endpoint to instantaneous EEC is acceptable as a conservative Tier I approach; however in Tier II it must be compared to a time-weighted average EEC. Currently this is rather open ended. The following is stated in the text:

“Normally the distribution of time-weighted average EEC’s will be used to correspond with the toxicity endpoint concerned; however the choice of using the maximum to time weighted average EEC’s can be made independently for each endpoint after considering the relationship between LC50 and ECx and exposure, the mechanism of action, and the information on related compounds.”

Comparison of chronic endpoints with peak exposure values represent a serious mismatch of toxicity and exposure data that can result in significant mischaracterization of potential risk, particularly for environmental short-lived compounds. If ECOFRAM is serious in supporting the current wording, then specific examples are needed on when, in Tier II, it would be appropriate to compare chronic endpoints to an acute EEC and what data generated in the current guideline studies would lead to such a conclusion. Since this is discussed as part of Tier II, additional effect studies designed to look at some of the factors alluded to in the verbiage quoted above will generally not be available.

If one assumes that the primary objective of ECOFRAM was to develop a RA system which improves our ability to perform probabilistic assessments, then clarification is needed on which of the many recommended changes truly addresses this objective versus simply expanding the information data base. Similarly, all studies (current and future) should be evaluated relative to contribution to probabilistic exposure assessment. It seems that a number of studies, both current guideline studies and recommended new studies, do not contributed greatly to our ability for probabilistic assessments. Examples include the current soil dissipation and aged leaching studies.

It is necessary for ECOFRAM (or someone), to quantitatively estimate the precision of the risk estimate produced by each Tier. The actual risk curves should be bounded by upper and lower confidence limits. The exercise of bounding the risk estimates by estimating the various sources of uncertainty is not necessary for every risk assessment, but each tier needs this done generically or for a case example. The information which may be gained is an estimate of the precision of each method, which is important for decision making.

Chapter 3: Aquatic Exposure:

Overall this chapter is well written and informative; however it is very difficult to determine the bottom -line. The number of recommendations is overwhelming!! There are so many recommendations that it is unclear how close or far we are from realizing the different recommendations. First, it seems that much of what is discussed is technically doable and in many cases has simply not been used. A listing is needed of exactly what is currently possible. Then the question needs to be asked, if it is possible, why isn't it happening??

A primary objective of ECOFRAM is to develop probabilistic tools. It is unclear as to what aspects of the exposure and effects analysis actually are probabilistic. Many of the inputs values to models appear to be deterministic and models such as EXAMS are themselves deterministic. This will affect how outputs (and risk) can be interpreted. An example of the different types of exposure curves and how they are to be interpreted would be very useful

The current Tier II modeling (worst case site PRZM / EXAMS runs) needs to remain as an option. The proposed tier system goes from a simple deterministic approach (GENEEC) in Tier I to a probabilistic approach in Tier II which addresses spatial and temporal factors. It seems there should be room for an intermediate tier in which a deterministic model (PRZM, EPIC, etc.) could be used to show how a particularly important E-fate pathway for a compound can influence its environmental fate. For example, suppose a compound is applied to a closed plant canopy and has very low wash off potential. In this case, the compound is not a risk to surface water (spray drift excluded), but if it fails Tier I, it will undergo relatively extensive probabilistic modeling. It seems appropriate to have an opportunity to describe the limited runoff potential by configuring a model to properly resolve the limited wash off at one to a few sites to demonstrate that this unique property is the basis of limited exposure. It seems this opportunity is lost by a quick transition into an "automated" probabilistic process. If this cannot be properly captured in the automated Tier II, then resources will be wasted as it moves into Tier III.

Current Tier I (GENEEC) is based on old (PRZM 2) modeling. It should be updated to reflect PRZM 3.12 if it can be expected to provide relative (ca: 95th percentile) results. Note: there was a suggestion to update GENEEC in section 3. This comment should be made in section 2 and made more clearly in section 3.

Although there is discussion of marine / estuarine exposure, there is little indication on how such exposure estimates are to be performed. What model and scenario is used to estimate marine organism exposure?

Soil Dissipation Studies: Efforts to harmonize soil dissipation study guidelines between Canada and the US should reflect the needs outlined by ECOFRAM; although presently the draft of the harmonized soil dissipation guidelines does not reflect any consideration of ECOFRAM needs nor recommendations.

An increased scope for soil dissipation studies is contained in the draft ECOFRAM Exposure

Assessment chapter. These recommendations include investigating “foliar processes” and “insects and soil invertebrates” and “air sampling” (Exposure Assessment, page 3-94 & 3-97) as part of the soil dissipation study. This is inconsistent with EPA current guidance that soil dissipation studies should be “bare ground” studies in order to maximize leaching. This “maximized leaching” concept suggests that soil dissipation studies should be used as triggers for ground water monitoring studies, however, present laboratory data (e.g., half-lives and Koc values) serve as the principal triggers for ground water monitoring studies even when soil dissipation data do not indicate ground water studies are justified.

“Validating mitigation effectiveness” is an unclear concept?? The need for mitigation is likely based on unvalidated modeling, so we really do not know the accuracy of the original risk estimates. Are we wanting to validate mitigation effectiveness or validate our risk assessment decision? It will be very difficult to validate a risk assessment decision since it is assumed, and rightly so, that any decision made should result in minimal environmental risk (especially for new products). At least, the potential effects would be minor enough to make actual field measurement of such effects improbable. If we want to evaluate the effectiveness of any given mitigation approach in reducing the exposure potential, it is inefficient to approach this on a chemical by chemical basis. A government / industry Task Force approach would be much more effective in the development of a cohesive data base on the effectiveness of various mitigation practices.

References to FQPA monitoring studies and tying these to Ecological exposure monitoring is inappropriate and outside scope of workgroup. The scope and objective of such studies are quite different and it is difficult to understand the purpose of including such statements in this report. Was trying to connect drinking water and ecological exposure monitoring studies really discussed and agreed to by ECOFRAM as a useful and doable objective?

The report notes that ECOFRAM strongly recommends that *“widespread monitoring be carefully utilized and that the results from monitoring studies not be given undue emphasis. Unlike predictive modeling they only represent one scenario in one season and can prove misleading...”*

This statement is too strong and gives the impression that models are “right” and real world data is questionable. Later in the same paragraph it is stated that *“When used in concert, modeling and monitoring can make a powerful combination...”* This is the key concept to be stressed in this paragraph - not that models are “right” and real world data is questionable.

The term “widespread monitoring” needs to be defined. Since monitoring is not “triggered” until Tier IV, it would seem that any monitoring to be conducted would be focused and fairly site specific as opposed to “widespread”. Most of the advanced modeling efforts should have identified the specific conditions under which unacceptable exposure may occur and any monitoring efforts should be designed around such information. This may - or may not - be “widespread”.

Benchmark / Regression / empirical modeling represents another approach which has potential value but which was virtually ignored by ECOFRAM. The advantages and disadvantages of such an empirical approach should be addressed. There is a tremendous amount of monitoring data currently being generated via a variety of government /academic and industry efforts. It is possible that the use of these large data bases to develop empirical models is more practical than trying to validate mechanistic models, especially at the watershed level where it is extremely difficult to include all potential mitigating factors. Ignoring this approach is simply unacceptable.

Although the report discusses uncertainty, what is still needed is a description of how expressions of uncertainty should be used in the risk assessment and decision making process. The first page and half of the aquatic effects chapter (4.1) gives a excellent overview of the types of uncertainty and the limitations of some of the probabilistic methods (e.g., Monte Carlo). These concepts should be kept in mind when reviewing all sections of the document since they provide a reality check on what is being achieved by the various recommendations made by ECOFRAM.

Describing the model used and all types of uncertainty associated with the various Tiers (or Levels of Refinement) of the risk model is critical in justifying and defending why a higher tier assessment is “better” than a lower Tier assessment. It also is important in making risk management decisions by providing a level of comfort to the risk manager. Describing the overall risk assessment model is critical in assuring the transparency of the risk assessment (bias and conservatism should be clearly apparent).

A practical consideration is that often as a model complexity increases, the “precision” of the risk estimate can apparently decrease due to increase parameter error related to the difficulties in estimating additional parameters. As one moves through the Tiers, model error would be expected to decrease, but quantified uncertainty associated with parameter estimation and natural stochastic processes increase (of course they likely do not increase since they were simply not considered at the lower levels or where covered by conservative assumptions.). Given this, it would seem possible that uncertainty would actually increase as one moves through the Tier process. Is this possible?

A concern has been expressed by some over the use of MUSCRAT. Conceptually this model captures the essence of a temporal / spatial standards scenario type model and it is a good first attempt at getting to a probabilistic exposure value for ecorisk assessment. However, it does not address uncertainty in PRZM/EXAMS parameterization and may lead to overly conservative estimates of exposure (90th percentile of 90th percentile predictions - see discussion below). Sufficient computer power and sensitivity analysis/Monte Carlo tools now exist to deal with uncertainty around individual parameters in the model. These concepts should be built into a new probabilistic model which includes many of the concepts in MUSCRAT but allows for a more critical assessment of uncertainty. It is important to note that this tool must be relatively “hard wired” with respect to scenario development (much like muscrat) to allow efficient use by regulators (see final note below)

Pages 3-42-46 provide an excellent description of MUSCRAT. Section 3.4.4.3.2 provides a list of assumptions and limitations. Additional limitations which again add to the point that a new model is necessary for Tier II include:

- The selection of soils to represent bins, and the parameterization of the soils in PRZM needs to be transparent (published if possible). An algorithm was devised to extract the soil nearest the centroid of each bin. When this soil was problematic (perhaps it had weathered bedrock at some depth to prevent groundwater modeling), another soil near the centroid was chosen. This procedure and the parameterization of each PRZM file for the specific soil should be considered.
- The MUSCRAT concept affords the use of multiple soils, but it does not address sensitivity in parameter estimation. Each soil is still described by single values (by horizon) for critical parameters, such as:
 - Half-life
 - Kd
 - Field capacity

Thus MUSCRAT provides 25 “representative values” with the same level of uncertainty as any single site analysis. A general lack of information regarding the specific values used to represent each soil (refer to item 1) makes it difficult to have confidence in the results.

- MUSCRAT provides the upper 10th percentile exposure concentrations for each bin, which is area weighted for the specific crop of interest. The results (see figure 3.3) provided by MUSCRAT is an exceedence probability of these 25 results. Thus, if one looks at the 0.1 value for use area exceedence, it is essentially the upper 90th percentile of upper 90th percentile predictions from 25 representative soils (area weighted). A rigorous statistical interpretation of this with respect to occurrence in the region of interest is required for MUSCRAT and should be included in any future replacement product.

Chapter 4: Aquatic Effects

General Comments on Effects Chapter:

The effects chapter comes across as being separate and issue-based. Additional text is required to integrate specific sections and relate them back to the probabilistic risk assessment objective. Examples of risk assessments incorporating different exposure output, population models, TTE analysis models and both acute and chronic endpoints may help the integration.

Triggers for additional chronic study are very unclear and open-ended? Who determines and when / how?? The time for effects to occur analyses also highlights a problem with current subchronic and chronic guideline test designs. There is little mentioned about revising these particular testing protocols; however for these tests to be really useful for risk assessment

purposes, the type of data collected, and when these data are collected, needs to be modified to provide useful inputs for risk assessment. The report calls for additional chronic testing, but it is unclear whether this means more of the same (same type of tests, do more fish or crustaceans), or whether the report is recommending that the suite of organisms evaluated be expanded to include representatives from more phyla/niches.

Vulnerable Headwater Streams” are identified in Tier I as the ecosystem being assessed. Additional explanation is needed to explain this versus the Tier I farm pond.

Use of full dose response curve for analysis of acute data and a regression approach for chronic data is good. The criticisms outlined in the report on the problems associated with the use of the NOEC as the regulatory endpoint in chronic studies are appropriate. However the “x” level to be regulated on must be defined beforehand and be consistent. The experimental design can be influenced by the confidence one desires in a specific “x” level. Also, it needs to be realized that it can be quite difficult to obtain chronic dose response curves; obtaining such curves should not become a requirement of the study. One needs to be sure there is enough flexibility in the guidelines so that a dose-response curve does not become the criterion for judging a chronic study. For many risk assessments, a simple NOEC will be adequate.

Population Models:

Relative to population models, quantitatively predicting the effects on populations is a noteworthy goal, but does it fit into a pesticide regulatory framework? If there is a specific resource of concern then such models can be a helpful tool, but if there is a general concern about fish or zooplankton population, where does the requirements for toxicity data to support the model stop? Also, obtaining the relevant age specific mortality and reproduction data can be difficult. If one decides to use population models in the broader generic sense one need's to come up with a “worst case” scenario (species life history, etc.) for population modeling. With the number of species available to choose from, it is likely that one could always find a theoretical population with a low “r” and high sensitivity to justify that a compound is having population level impacts. Given this, what is to be gained?

It would be very helpful to have an example of how exactly a population model would be used such that the additional information generated by the model would allow a regulatory decision to be made that could not have been made at the Tier III level.

Joint Probability Distribution (JPD):

Given statements made earlier about the tremendous complexity associated with all of the recommended changes in the higher Tiers, it would be useful if ECOFRAM would recommend just one type of JPD at Tier 2 instead of a choice (page 4-21 “.... concentration effect curves for acute mortality, time to population recovery, or other effect endpoints).

Confidence bounds on dose response relationship should be reported and, if possible, these bounds should be propagated into bounds for the JPD.

A Joint Probability Curve based on a single species dose response is fine in theory, but may be difficult to implement. Does one use the probit model or raw data. If raw data is used, a typical maximum of 5 points could be available to draw the curve. In how many of our data sets do we have 5 different responses? If the model curve is used, it must be remembered that the probit model does not do a good job at estimating 0 and 100% kill! Choice of dose-response model can be important (Probit, logistic, etc.) Current dose-response models are chosen for their accuracy in estimating LC50, not the full dose-response relationship.

JPD's for single species versus species sensitivity seem to address different assessment endpoints. Is the assessment endpoint a certain percentage of a given population or are we addressing aquatic communities and protecting a certain percentage of the species present - or both?. Is it logical (other than the fact that additional data may be needed for the species sensitivity assessment) that these JPD's are in different Tiers? I assume the assessment endpoints do not change as one moves through the Tiers? Using single species JPD's, is it assumed that if you are protective of a population of *Daphnia* (surrogate species) based on Tier I/II evaluation, that you will be protective of the community? To follow, when one has JPD's for single species and for species sensitivity, which one is used for the final regulatory decision? If the JPD for species sensitivity indicates less risk than that based on the single species assessment, will this be the decision or will EPA revert back to the single species JPD for their regulatory decision. If EPA reverts back to the single species JPD, why would industry want to spend additional resources evaluating the species sensitivity distribution?

Time-Varying or Repeated Exposures:

Pulse-dose experiments represent a very useful tool that can help to reflect more realistic exposure scenarios. For some short lived compounds these studies can greatly increase our ability to more accurately reflect the hazard under the defined exposure scenario. Of course the key point is whether the results will actually be used in place of the results from the constant exposure standard guideline studies. EPA must indicate a willingness to regulate on the results.

Time-to-event analysis:

On theoretical basis, Time-to-Event Analysis as an improvement in how toxicity data is reported. It can provide the basic time-concentration-effect model which controls the biological response. However, it is possible that the claims of usefulness in the regulatory context are a bit exaggerated or impractical as summarized in the ECOFRAM chapter. In addition, in order to do a good job with this concept, additional test levels and observations periods may be required to be added to the current standard toxicity tests (depending upon the compound). It is unclear, from a regulatory decision making perspective, whether the extra effort will make a difference in the conclusions to be drawn concerning the acceptability of a given level of risk. In the case

studies hopefully to be performed, the usefulness of having such data in making regulatory decisions must be decided upon.

Sediment Toxicity Tests:

I found this section to be a bit confusing, both technically and from an ECOFRAM process perspective. This chapter appears to be written by EPA, as an EPA overview of their current position on sediment toxicity tests. Is this how the ECOFRAM process worked? Was there input and discussion among the scientists on the ECOFRAM group on the positions and topics presented in this Section? If not, this Section should be removed from the final Chapter since it does not reflect the opinion and position of the group but rather that of EPA.

Specific areas to be discussed include the following:

The report should be revised (page 4-108) to reflect that OPP will not be calculating Sediment Quality Criteria, but will be relying on the basic EqP approach. It is very important that ECOFRAMs support of EqP is clarified and that the sediment section does not confuse the situation.

Sediment Testing Triggers (Page 4-109) - This section provides little clear guidance! ECOFRAM should recommend specifically what they (all members) believe are the appropriate triggers. The trigger/criterion "...concentration in interstitial water is equivalent to concentrations known to be toxic in the water column" is redundant with the other triggers noted since pore concentrations will come from PRZM/EXAMS, where the Koc controls the pore water concentration! I presume the point is that if you exceed an invertebrate water column LOC, then one should conduct sediment toxicity tests? What data (Guideline study requirement?) is to be used to evaluate the triggers is unclear. Clear triggers based on available data are needed!

The European sediment testing triggers should be supported. As a minimum EPA should use the same parameters, adsorption, toxicity, and persistence, and similar studies to determine these, even if the actual trigger values are different.

Any test guidelines should avoid the need to measure pore water in routine sediment toxicity tests. Pore water measurements are difficult, expensive and highly unreliable. Spiking methods and pore water recovery techniques are critical methods that need to be developed and defined in a regulatory context before measured pore water concentrations will have any usefulness. EqP and a properly measured OC estimate in conjunction with measured bulk concentrations should be used to estimate pore water values. This is consistent with how a EXAMS exposure estimates are generated.

The discussion on page 4-110 (lines 19-29) is extremely confusing. I think the conclusion is that bulk sediment concentrations will be used for the EEC in the RQ determinations. This is inconsistent with the pore water focus mentioned earlier in the document (see below). Also any

use of bulk concentrations must discuss that the OC content must be the same for the exposure and effects analysis

I have difficulty with the concept of having to use marine species in addition to freshwater species? Both are simply surrogate species and, given the large distribution in sensitivity for all organisms, it is likely irrelevant whether they are marine or freshwater. If we are using marine organisms to represent the marine environment, how are the sediment concentrations in the marine environment going to be estimated?

Earlier in the document (pg 4-27: 4.2.3.5), ECOFRAM noted a number of key points relative to sediment testing which I agree with, but which are somewhat contradictory to statements made on pages 4-111. Specifically:

Page 4-27, lines 29 - 30: "If sediment toxicity is found to differ substantially from that expected based on pore water concentrations, tests with additional sediment types may be required."

Page 4-27, lines 32-33: "If the sensitivity of the benthic species is found to be comparable to that of the pelagic species already tested, data for pelagic species can be used to estimate the distribution of species sensitivity for benthic organisms"

The document needs to be modified to correct this inconsistency.

As an alternative perspective on sediment toxicity tests, it is possible that something has been missed concerning why sediment tests should be performed and how they should be used. There is really no reason to think that as a group sediment dwelling organisms are more sensitive than pelagic species. The only difference is that sediment organisms have the potential to be exposed to a chemical via additional exposure routes - particle ingestion and contact. For most chemicals and species the toxicity of a chemical is primarily correlated to the pore water concentration, not the bulk (or particle) concentration. For most benthic species a toxicity value from a water column test will be predictive of an organism response in the sediment. A sediment pore water LC50 will in general be equal to the organism water column LC50.

Given this, a primary reason to conduct sediment toxicity testing for pesticide regulation is to test the hypothesis that non-water routes of exposure are unimportant. To evaluate this hypothesis a water column and sediment toxicity need to be conducted with the same species. If non-water routes are important then additional testing may be needed. If the theory holds, as it appears to in most cases, then only the water column RA (with benthic species included) is needed.

Microcosm / mesocosm Studies:

The discussion on the use of microcosm / mesocosms provides a fairly accurate overview of the potential scientific usefulness of such studies. Designed correctly, they can provide a wealth of information on the potential primary and secondary effects of a pesticide under the designed test conditions. However, I would question the following statement made on page 4-121 (lines 6-8):

“... most experts (and ECOFRAM) now believe that microcosms and mesocosms can provide valuable information for higher-tier ecological risk assessments if each study is customized to address the specific concerns for a particular chemical.”

Relative to this statement, I question whether the problems associated with the past use of the results from such studies in a regulatory context have been fully considered. A number of issues exist relative to the usefulness of such studies in addressing critical regulatory issues. Specifically:

1. Will the results from mesocosms / microcosms reduce uncertainty and allow for better decisions?

Microcosm / mesocosms can certainly be provide additional information; however I do not believe they will significantly reduce the uncertainty in the risk assessment. Such studies represent a “snap shot” of potential effects in static systems which result the specific exposure scenario chosen. Given the vast array of potential exposure scenarios and ecosystems, I do not think uncertainty can be greatly reduced by such studies and certainly not reduced enough to justify the extensive costs (to both industry and EPA).

2. Will the results improve predictability of potential adverse effects actually occurring in the field?

For reasons similar to those outlined above, I do not believe that predictability of effects will greatly improve. Of all the studies I have been involved with over the last 10 years, none of them have provided any additional insight or improved one’s ability to **predict whether unacceptable adverse effects may actually occur in the field.**

3. Are the results useful in probabilistic assessments?

As noted above, extensive additional data is generated; however none of the data will be useful in improving probabilistic risk assessments. As noted in the chapter, one of the useful features of mesocosms is providing toxicity data on additional species. I agree that this is useful; however I think it is much more useful, and economical, to generate such data in laboratories.

4. Is the value obtain relative to an improved risk assessment worth the cost of study conduct and resources for study review and interpretation?

In terms of reducing uncertainty and / or improving predictability I do not think these studies provide value that is justified given the costs. However given the attractive nature of having such studies, especially for biologists / ecologists interested in whole ecosystem level assessments, there may exist the desire to frequently request such studies. I would strongly request that in such cases the risk managers critically review such requests and truly establish that the results to be obtained from such studies will greatly improve their ability to make a final decision.

Detailed, page specific comments:

Page	Comment
2-7 line 28	The risk assessor should characterize the uncertainty but the risk manager should make the decision as to whether it is acceptable for decision making.
2-9 line 17	Substitute “efficacy” with “effectiveness”
2-11 line 2, Fig. 2-4	The risk assessor should characterize the uncertainty and the assumptions made. The risk manager should determine whether the assessments (including uncertainty and assumptions) is adequate for decision making.
2-12 line 22-24	The report indicates that EFED has the final say on risk assessments and risk management decisions. The latter decisions are currently handled by the Registration or Reregistration and Special Review Divisions, and should remain with those divisions.
2-18	Tier I is intended to be relative to Tier II rather than the real world. This is based on the desire to approach 95 th percentile of Tier II results. The conservative aspects of Tier II must be more clearly defined in this case.
2-21, line 12	Tier II objective should not be to confirm that risk associated with Tier I “still applies” based on the more detailed analysis! It should be to further evaluate the risk for the assessment endpoints not eliminated as “minimal risk” in Tier I.
2-25	The application of joint probability curves does not reflect the conservative nature of the Tier II exposure assessment. Line 23 on page 2-29 makes this statement, but it should be made sooner in the development of the interpretation. It must be clearly explained that Tier II still represents a compilation of worst case assumptions> These should be clearly spelled out.
2-25 Fig. 2-7	Should discuss whether this risk assessment would be the same for chronic and for endpoints other than mortality.
2-31 line 3	Are wetlands and estuaries only addressed at Tier 3?
2-31 line 7, 8	This infers that Tier 3 AgDRIFT only applies to Tier 3 risk assessments. Is this correct ?
2-34 Fig 2-11	To make this figure comparable to Tier 2, the JPC should be shown.
2-37 line 19+ .	DO NOT RELATE FQPA WATER MONITORING TO ECOTOX MONITORING. The scale is very different (water sources large enough to serve as reliable DW sources vs. edge-of-field ponds and/or streams). The overall purpose and approach to such studies can be quite different and they should not try to be combined. All discussion / reference to FQPA should be deleted from the ECOFRAM document.
2-37 line 26	Unclear what is meant by “carefully utilized”
2-40 line 4-6	Need to make it clear that the concentration that you are referring to is in the abiotic compartment and not the organisms otherwise the concept of

	bioavailable could become very confusing (i.e. can be amount adsorbed from gut).
Section 3.1	An overview or diagram of the various exposure models discussed and how they link into the exposure assessment would help the reader follow this section (PRZM, EXAMS, GENEEC, AgDRIFT, RADAR, MUSCRAT etc). Need to indicate which models are currently in use and which are still in development/being implemented.
3-1 line 23	An example of a “vulnerable headwater environment” would be beneficial. It is unclear to me how Tier I assessments are addressing headwater streams?
3-3, line 26	I should support the concept that “Studies that are not used at all in risk assessments should be made optional.... Are Terrestrial Dissipation Studies an example of such a study. If they are not used in the assessment, are they of value?
3-9 line 6-7	Is this referring to GENEEC output ?
3-10, line 29+	ECOFRAM should more clearly define how and when it encourages MUSCRAT to be used. For example, in Table 2-3 (pg 2-42) the following is provided: “MUSCRAT 2 (PRZM-EXAMS)”. Does this suggest a new version of MUSCRAT, or the use of PRZM-EXAMS. On page 3-11, line 11 it is suggested that MUSCRAT be embellished to support Tier III issues. Is ECOFRAM recommending MUSCRAT as a tool for future use, or are they recommending that the conceptual foundation around which MUSCRAT was developed should be used to develop and new model for Tier II?
3-11 line 4	Sentence incomplete
Section 3.3	This section does not belong in exposure but in a chapter common to exposure, effects and risk.
Section 3.4 pages 16-25	This information would be best served in an appendix because it detracts from the proposed exposure assessment.
3-51 line 14	According to this table, GENEEC only deals with spray drift and not spray drift and runoff. Needs to be corrected.
3-61,65 Table 3-5, 3-6	Which of these factors can have inputs as distributions ? Are all of these inputs single values ? Will this be changed as part of the probabilistic exposure assessment? Section 3.6.3.1 (pg 84) discusses selecting single input values. Isn't the ideal situation to input a distribution to reflect the actual data or the uncertainty around a single value?
3-73+	Recommendation to add study types or intensity should be based on the need to improve predictions of chemical fate in the fate/transport models. ECOFRAM should perform a sensitivity analysis to determine if the substantial increase in chemistry studies will result in substantially improved predictions for Tiers I-III (some sort of cost/benefit and sensitivity analysis should be performed prior to increasing registration demands)
3-90	I agree that time varying exposure should be considered, but it seems like a difficult process to derive toxicity data to support the infinite possibilities of

	transient exposure. Definitive guidance will certainly be needed to make such studies possible in the regulatory area?
3-124 line 20-24	It seems unjust to critique AgDRIFT on the basis that it is deterministic and prescribes buffers since 1) It is only a component of the exposure assessment 2) it is very unclear what is deterministic about the runoff models other than rain storms (most of the inputs appear to be a series of conservative deterministic inputs) 3. current runoff models cannot quantify effect of buffers.
4-1,2	The definition of uncertainty types is good. The exposure section should use the same definitions
4-30 line 4-6	There is a good reason why environmental regulations focus on exposure intensity rather than duration since options for mitigating exposure will primarily affect intensity. Therefore it can be argued that focusing on duration is not equally appropriate.
4-109 line 19-33	Need to state what ECOFRAM recommends (i.e., do they support Maund et al. Triggers). The text focuses on OPPTS recommendations which are not directly relevant.
4-113 line 6-8	It is not very evident how endpoints for endocrine disruption can be “easily” incorporated to the ECOFRAM risk assessment tiers. This is because it is unclear how other non-lethal endpoints from chronic tests are to be used in the risk assessment. Use of NOEL in risk assessment or JPC ?